

OVERVIEW OF THE 1997 DARPA SPEECH RECOGNITION WORKSHOP

David S. Pallett, Workshop Chair

National Institute of Standards and Technology (NIST)
Information Technology Laboratory (ITL)
Room A 216, Building 225 (Technology)
Gaithersburg, MD 20899
E-mail: dpallett@nist.gov

1. THE WORKSHOP FOCUS

This year's DARPA Speech Recognition Workshop focussed on one multi-faceted technical challenge—recognizing the "found speech" in radio and television news broadcasts. As such, this challenge is an expansion of last year's Hub 4 Broadcast News Benchmark Tests.

This year's Hub 4 Broadcast News tests made trial use of a new paradigm—a "Partitioned Evaluation" (PE), which was intended to contrast with an "Unpartitioned Evaluation" (UE), and which made use of information derived from extensively annotated and time-marked reference transcriptions which in some cases obviated the need for "chopping" or segmentation modules as components in the Hub 4 systems. The annotation convention was intended to "partition" the data into a number of "focus conditions" with similar attributes—"baseline" speech (F0) typically including read or "planned" speech collected in studio conditions, "spontaneous" speech (F1), typically containing evidence of disfluencies, but also typically collected in studio conditions, "telephone channel" speech (F2), etc. More precise definitions of the attributes of each of the focus conditions can be found in other papers in this Proceedings.

In contrast, the "Unpartitioned Evaluation" paradigm permitted essentially no use of information derived from "side information". Three sites that had participated in last year's Hub 4 tests (BBN, CMU, and IBM) provided results for both PE and UE tests.

This year, no Hub 3 tests were reported upon at the Workshop. Other related work involving conversational speech over telephone channels, sometimes referred to as "Hub 5" (which has been the focus of other workshops) was reported upon at a session devoted to "Other Relevant Research".

2. THE WORKSHOP SESSIONS

Not surprisingly, the Opening Session of the Workshop, on Sunday Evening, February 2nd, included the now-traditional reviews of properties of the Hub 4 corpora and test materials,

discussion of the test paradigm, and presentation of the "official" NIST summaries of the test results.

The initial presentations on Monday were devoted to "High Level System Overviews". These were followed by two Technical Sessions and an informal evening Session which provided opportunities to explore and listen to properties of the Hub 4 test data, and to discuss potential changes to the test paradigm.

The Technical Sessions continued on Tuesday, with some "free time" on Tuesday afternoon, followed by an informative Technology Demonstration Session, and an evening session devoted to discussion of the planned "Spoken Document Retrieval" task to be included in next year's Text Retrieval Conference (TREC), organized by Donna Harman and her colleagues at NIST.

Sessions on Wednesday included one on "Other Relevant Research", and two somewhat informal ones—one on "Future Plans" and a "Wrap-up and Final Discussion" Session.

3. DIFFERENCES BETWEEN THE WORKSHOP "AGENDA" AND THIS "WORKSHOP PROCEEDINGS"

While the structure of this Proceedings largely follows the sequence of oral presentations at the Workshop, there are some noteworthy differences. The informal nature of some sessions (e.g., the Tuesday evening session devoted to demonstrating properties of the test data, the Technology Demonstration Session, and the closing sessions discussing Future Plans) did not yield papers for inclusion in the Proceedings; these materials were generally incorporated into a single paper.

It's also noteworthy that this is the first Workshop in this series to produce Proceedings in several media: (1) traditional paper, (2) CD-ROM, and (3) a "web-accessible" on-line version. From the web and CD-ROM versions, papers can be accessed in several formats: printable Postscript files, Acrobat Portable Document Format ("PDF") files and, in most cases, files in HTML (some of which have playable audio files).

4. THANKS

As noted in a similarly-titled note in the Proceedings of the February 18-21, 1996 Speech Recognition Workshop sponsored by DARPA, "over the past decade many individuals have been fortunate to participate in... DARPA Workshops... and these individuals have helped advance the state-of-the-art of automatic speech recognition through their own research". This year, approximately 140 individuals participated in the Workshop.

I agreed to serve as Workshop Chair, and was again fortunate to have the help of an Organizing Committee that included Patti Price, Roni Rosenfeld, Salim Roukos, Richard Schwartz, Richard Stern, and Phil Woodland. Our business was efficiently completed using E-mail and conference calls. Rich Stern also deserves credit for Chairing the Working Group that developed the specifications for the 1996 Hub 4 Tests.

Special thanks are due to the participants in the Technology Demonstration Session, especially Charles Hemphill for his presentation of TI's "Speech-Aware Multimedia for Novice and Expert Users", Sean Colbath for demonstrating BBN's "Speak'n'Surf: Speech Access to Information", and Dave Stallard for BBN's "SPIN: SPeech over the INternet". Their willingness to run the risks of real-time system crashes in front of a large and technically sophisticated audience, and never-failing good humor are greatly appreciated!

Sharon Kaufmann once again made everything happen - especially in making arrangements for us to meet in a new venue for our community - the Westfields Conference Center, in Chantilly, Virginia, conveniently located near Dulles International Airport. This year's special thanks are for her patience in finding a creative and expeditious way to photocopy many more papers for the Notebooks than had been anticipated. Our community owes a huge debt of gratitude to Sharon for making lots of things happen, and, in some cases, for us to further our research and pay our bills!

At CMU, Carol Patterson helped prepare Notebooks for distribution at the Workshop. At NIST, Bruce Lund agreed to serve as Editor of our first Proceedings in CD-ROM and web-accessible format, as well as the traditional printed form.

Although last year's "mountain-top" experience at Arden House was hard to follow, it's hard to fault the comfort and productive ambience of Westfields. Both venues offer fine meals and comfortable accommodations, but having a tuxedo-clad pianist serenade us on a Boesendorfer grand piano before dinner was an unusual and unprecedented pleasure.